

## KLASIFIKASI JENIS MANGGA APEL MENGGUNAKAN METODE K-MEANS KLUSTERING

Agyztia Premana<sup>1</sup>, Otong Saeful Bachri<sup>2</sup>, Akhmad Pandhu Wijaya<sup>3</sup>

<sup>1,2</sup>Universitas Muhadi Setiabudi, <sup>3</sup>Universitas Wahid Hasyim  
a.premana@umus.ac.id<sup>1</sup>, otongsaefulbahriumus@gmail.com<sup>2</sup>,  
pandhudsn@unwahas.ac.id<sup>3</sup>

Received : 28-09-2022

Accepted : 20-10-2022

Published : 25-10-2022

### Abstract

**Introduction:** The rapid development of technology is pushing it further, making people more comfortable in many fields, including industry. Mango can be processed into various types of food. Using sweets, various processed products, and various mangoes. One example of the current impact of technology on the industrial sector is the potential for systems to self-study (automatically) like humans. This is a process known as an artificial neural network. **Purpose:** Knowing and analyzing the classification of mango species using the K-Means Clustering method. **Methods:** This study uses the K-means clustering method, in which the system is built by applying an artificial neural network to the modeling and extraction of RGB values and standard RGB matrices, circumference, area, length, width, shape, and slenderness. **Results:** Based on the experimental results, the computation time for the Mango image required for the feature extraction process for each dataset is on average 0.85 seconds, and the computational time for training data on the test data is an average of 0.006 seconds. **Conclusion:** In this study, it can be concluded that the average computation time of mango image for each dataset is 0.856 seconds. The use of 1 hidden layer is more effective with the highest accuracy and the fastest time than using 2 hidden layers.

**Keywords:** Classification, Mango, Image Processing, Backpropagation Method.

### Abstrak

**Pendahuluan:** Perkembangan teknologi yang pesat mendorong lebih jauh, membuat masyarakat lebih nyaman di banyak bidang, termasuk industri. Mangga dapat diolah menjadi berbagai jenis makanan. Menggunakan manisan, aneka produk olahan, dan aneka mangga. Salah satu contoh dampak teknologi saat ini di sektor industri adalah potensi sistem untuk belajar sendiri (otomatis) seperti manusia. Ini adalah proses yang dikenal sebagai jaringan saraf tiruan. **Tujuan:** Mengetahui dan menganalisis klasifikasi jenis mangga dengan menggunakan metode K-Means Klustering. **Metode:** Penelitian ini menggunakan metode K-means klustering, pada sistem dibangun dengan menerapkan jaringan syaraf tiruan pada pemodelan dan ekstraksi nilai RGB dan matriks RGB standar, keliling, luas, panjang, lebar, bentuk, dan kelangsingan. **Hasil Pembahasan:** Berdasarkan hasil eksperimen, waktu komputasi citra Mango yang dibutuhkan untuk proses ekstraksi ciri untuk setiap dataset rata-rata adalah 0,85 detik, dan waktu komputasi untuk data latih pada data uji rata-rata 0,006 detik. **Kesimpulan:** Pada penelitian ini dapat disimpulkan bahwa rata-rata waktu komputasi citra mangga untuk setiap dataset adalah 0,856 detik. Dan penggunaan 1 hidden layer lebih efektif dengan akurasi tertinggi dan waktu tercepat dari pada menggunakan 2 hidden layer.

**Kata kunci:** Klasifikasi, Mangga, Image Processing, Metode Backpropagation.

Corresponding Author; Otong Saeful Bachri  
E-mail: otongsaefulbahriumus@gmail.com



### Pendahuluan

Perkembangan teknologi membangkitkan keinginan untuk meningkatkan efektivitas dan efisiensi di berbagai sektor, termasuk pertanian. Kemajuan teknologi dalam dunia pertanian sangat dibutuhkan untuk menunjang kegiatan yang ada di sektor pertanian, salah satunya adalah pengolahan hasil pertanian dan perkebunan. Industri pengolahan hasil pertanian dan perkebunan saat ini berkembang pesat,

terutama dengan berkembangnya teknologi produksi mangga yang ada (Sari et al., 2016). Di Indonesia sendiri buah mangga digemari oleh semua kalangan karena dikonsumsi langsung sebagai buah atau diolah menjadi produk konsumen lainnya seperti roti mangga dan selai mangga. Menurut Badan Pusat Statistik dan Administrasi Umum Hortikultura, produksi mangga terus berkembang selama tiga tahun terakhir. Mangga banyak diproduksi di Indonesia, namun pengolahan mangga tidak seimbang. Penelitian ini berupaya untuk mendukung pengolahan mangga tahan manga yang jumlahnya masih sedikit (Ristian, 2017). Ada beberapa tahapan pengolahan mangga sebelum produk yang dihasilkan didistribusikan ke konsumen. Salah satu tahapannya adalah memilih usahatani dan perkebunan mangga sesuai dengan kebutuhan pengolahannya.

Pemilihan varietas mangga sangat penting karena varietas mangga yang berbeda memiliki metode pengolahan dan hasil yang berbeda. Sortasi buah mangga dilakukan berdasarkan bentuk, ukuran dan warna buah mangga. Pemilihan warna meliputi rata-rata RGB dan standar deviasi RGB, ukuran meliputi keliling, luas dan lebar, dan bentuk meliputi kebulatan dan kelangsingan. Tujuan dari proses sortasi buah mangga ini adalah untuk menyortir buah mangga berdasarkan jenisnya. Petani atau penyortir mangga umumnya mengidentifikasi varietas mangga berdasarkan ukuran dan warna kulitnya. Ini adalah cara termudah (Malang, 2016).

Permasalahannya, meski sederhana, proses pemilihan varietas buah mangga bisa jadi tidak maksimal, apalagi jika dilakukan secara manual. Hal ini dapat terjadi karena adanya perbedaan persepsi masyarakat terhadap pemilih buah mangga dalam kaitannya dengan komponen warna objek, bentuk dan ukuran citra, meskipun objek yang dilihat sama persis. Apalagi dengan meningkatnya produksi mangga di Indonesia, seharusnya semakin banyak perusahaan dan pabrik yang mengolah mangga tersebut. Pabrik dan perusahaan pasti akan memproses mangga dalam skala yang sangat besar. Jika sebuah pabrik atau perusahaan mempekerjakan pekerja untuk memilih varietas mangga, pabrik akan memproses varietas mangga secara manual hanya untuk memperhatikannya. Kita harus mempekerjakan lebih banyak pekerja (Utami et al., 2017).

Berdasarkan permasalahan di atas, maka diperlukan penelitian untuk mengimplementasikan aplikasi identifikasi varietas mangga menggunakan pengolahan citra digital (*image processing*). Penelitian ini dilakukan karena dapat menjadi teknik alternatif untuk memitigasi atau mengatasi permasalahan yang ada. Membangun aplikasi untuk mengidentifikasi varietas mangga memerlukan penggabungan beberapa ilmu seperti pengolahan citra digital (*image processing*) dan jaringan syaraf tiruan (*k-means clustering*). Pengolahan citra digital (*image processing*) adalah teknik pengolahan citra objek secara digital, setelah melalui proses pengolahan, diperoleh informasi yang dapat diolah oleh komputer untuk berbagai keperluan (Aulianida et al., 2019).

Hal ini nantinya dapat diterapkan saat mengolah citra digital bentuk, ukuran dan warna buah mangga dalam klasifikasi jenis mangga. Untuk mengelompokkan biji mangga dan mengidentifikasi objek citra digital mangga, kita perlu mendukung bidang keilmuan lain: jaringan syaraf tiruan (*k-means*). Secara umum *k-means clustering* adalah bidang ilmu yang mempelajari teknik pengolahan data sehingga dapat digunakan untuk memberikan petunjuk-petunjuk tertentu yang bermanfaat. Metode *k-means*

memiliki banyak teknik dan metode untuk mengolah data, salah satunya adalah *backpropagations* (Nurmalasari & Pratama, 2018).

Diharapkan penelitian ini dapat memberikan solusi atas permasalahan yang ada pada pengolahan hasil pertanian dan perkebunan. Apalagi saat memilih manga apel berdasarkan kedewasaan. Penerapan aplikasi dalam penelitian ini dengan demikian akan membantu untuk mengolah hasil pertanian dan perkebunan secara cepat, tepat dan efisien sesuai kebutuhan konsumen.

### Metode Penelitian

Bagian metodologi menjelaskan langkah-langkah penelitian yang direncanakan. Termasuk cara memperoleh data, cara mengolah data, cara menganalisis data, cara menguji, dan terakhir membuat antarmuka pengguna alat uji untuk merancang dan membangun alat uji. Pengelompokan data dapat dibagi menjadi dua tujuan: pengelompokan untuk pemahaman dan pengelompokan untuk digunakan, Jika tujuannya adalah pemahaman, cluster yang terbentuk harus menangkap struktur alami data.

Proses clustering untuk tujuan ini biasanya hanya merupakan proses awal, setelah itu dilanjutkan ringkasan (rata-rata, simpangan baku), label kelas untuk setiap kelompok yang digunakan sebagai data latih klasifikasi. Jika tujuannya adalah penggunaan, tujuan utamanya biasanya untuk menemukan prototipe cluster yang paling representatif untuk data dan menyebarkan abstraksi dan setiap objek data ke cluster tempat data berada.

#### 1. Buah Mangga

Pengembangan halaman tampilan dapat ditambahkan untuk membuat grafik lebih mudah dibaca, memungkinkan pengguna untuk melihat akurasi yang dihasilkan dan bagian grafik waktu. Mangga (*Nephelium lappaceum*) merupakan buah asli Asia Tenggara seperti Indonesia. Setelah diperiksa lebih dekat, mangga menyerupai buah leci. Keduanya berwarna merah dan segar, tetapi mangga memiliki kulit berbulu sedangkan leci tidak berbulu, Selain itu, kandungan nutrisi dan vitamin buah mangga juga berbeda.

**Tabel 1. Kandungan Gizi dan Vitamin**

Nama	Kandungan
Protein	0,3g
air	86,6g
Lemak	0,1g
Karbohidrat	12,3g
Serat	1,6g
Kalsium	20mg
Fosfor	16mg
Zat Besi	0,1g

Kandungan nutrisi yang kaya dari mangga ini menawarkan berbagai manfaat kesehatan. Mangga merupakan jenis buah yang baik untuk sistem pencernaan. Ini karena sekitar setengah dari daging buah mangga mengandung serat makanan yang

tidak larut. Ini berarti jenis serat ini tidak bercampur dengan air dan langsung mengalir melalui sistem pencernaan Anda. Ini memudahkan pencernaan dan memudahkan mereka melewati kotoran saat buang air besar. Di sisi lain, serat larut dalam mangga menyediakan makanan bagi bakteri usus. Oleh karena itu, usus menghasilkan asam lemak rantai pendek seperti asetat, propionat dan butirat sebagai makanan untuk enterosit. (Singh & Misra, 2017).

## 2. Citra Biner

Citra biner adalah citra yang hanya memiliki dua nilai skala abu-abu, hitam dan putih. Meskipun citra berwarna sekarang lebih disukai, citra biner masih diperlukan untuk beberapa aplikasi, seperti pemrosesan citra untuk mengidentifikasi varietas mangga, seperti yang dibahas dalam penelitian ini. Gambar biner hanya memiliki dua warna abu-abu, hitam dan putih. Piksel objek memiliki nilai 1 dan piksel latar belakang memiliki nilai 0. Saat menampilkan gambar, 0 berwarna putih dan 1 berwarna hitam (Abdul & Adhi, 2013). Jadi dalam citra biner latar belakangnya berwarna putih dan objeknya berwarna hitam. Misalnya, gambar di bawah ini dengan ambang batas 0,7% (Sumaiya & Shantha Selva Kumari, 2017).



Gambar 1. Citra Buah Mangga.

## 3. Citra RGB

Citra berwarna atau biasa disebut dengan citra RGB adalah jenis citra yang warna-warnanya diwakili oleh komponen R (merah), G (hijau), dan B (biru). Setiap komponen warna menggunakan 8 bit (nilai berkisar dari 0 hingga 255). Oleh karena itu, kemungkinan warna yang dapat direpresentasikan mencapai  $255 \times 255 \times 255$  atau 16.581.375 piksel warna (Abdul dan Adhi, 2013). Table 2.4 menunjukkan contoh warna dan nilai R, G, B (Narendra et al., 2021).

**Tabel 2. Warna dan Nilai**

Warna	R	G	B
Hijau	255	0	0
Merah	0	255	0
Biru	0	0	255
Hitam	0	0	0
Putih	255	255	255

Model RGB biasanya direpresentasikan dalam bentuk kubus tiga dimensi, dengan warna merah, hijau, dan hitam pada dasarnya dan warna putih pada sumbu alur yang berlawanan. Fitur warna dapat diperoleh melalui perhitungan statistik seperti mean dan standar deviasi. Fitur-fitur ini dapat digunakan, misalnya, untuk tujuan identifikasi dan klasifikasi. Perhitungan dikenakan pada masing-masing komponen citra R, G, dan B (Shokrzade et al., 2021).

#### 4. K-Means Klustering

Metode pengelompokan dapat dibagi menjadi dua kelompok menurut strukturnya: pengelompokan hierarkis dan partisi. Pengelompokan hierarkis mencakup satu datum yang dapat dianggap sebagai grup, dua atau lebih grup yang lebih kecil yang dapat digabungkan menjadi satu grup yang lebih besar, dan semua data yang dapat digabungkan menjadi satu grup. Sampai saat itu, ada aturannya. Teknik pengelompokan hierarkis adalah satu-satunya teknik yang termasuk dalam kategori pengelompokan hierarkis. Metode partisi clustering membagi kumpulan data menjadi beberapa kelompok. Kelompok-kelompok ini tidak tumpang tindih antar kelompok. Artinya, setiap kelompok data hanya dimiliki oleh satu kelompok. Metode seperti K-Means dan DBSCAN termasuk dalam kategori pengelompokan partisi (Parida & Bhoi, 2017).

Metode pengelompokan secara struktural dapat dibagi menjadi dua jenis: hierarki dan partisi. Pengelompokan hierarkis mencakup satu datum yang dapat dianggap sebagai grup, dua atau lebih grup yang lebih kecil yang dapat digabungkan menjadi satu grup yang lebih besar, dan semua data yang dapat digabungkan menjadi satu grup. Sampai saat itu, ada aturannya. Metode pengelompokan hierarkis adalah satu-satunya metode yang termasuk dalam kategori pengelompokan hierarkis (Arslan & Arslan, 2021).

$$d_{ik} = \sqrt{\sum_{j=1}^m (X_{ij} - C_{ij})^2}$$

Clustering Metode partisi membagi kumpulan data menjadi beberapa grup tanpa tumpang tindih antar grup. Dalam artian formula diatas setiap kelompok data hanya dimiliki oleh satu kelompok. Teknik seperti K-Means Klustering dan DBSCAN termasuk dalam kategori pengelompokan partisi. (Alamuru & Jain, 2021).

#### 5. Fitur Ekstraksi

Fitur-fitur diatas dapat dibagi menjadi beberapa fitur lagi, jadi penjelsasan diatas merupakan beberapa fitur diantaranya.

### I. Kebulatan Fitur Bentuk

Kebulatan bentuk adalah perbandingan antara luas objek dan kuadrat Perimeter yang dinyatakan dengan rumus seperti berikut.

$$Kebulatan = 4\pi \frac{A(R)}{P^2(R)}$$

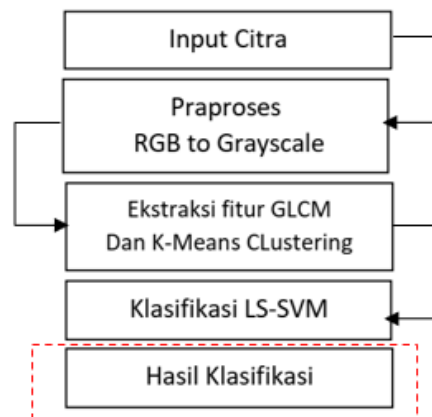
Dari formula diatas hasilnya berupa nilai citra  $\leq 1$ . Nilai 1 menyatakan objek R berbentuk lingkaran. Terkadang, future ini dinamakan kekompakan (Younesi & Amirani, 2017).

### II. Kerampingan Fitur Bentuk

Kerampingan bentuk adalah perbandingan antara lebar dan panjang yang dinyatakan dengan rumus sebagai berikut:

$$Kerampingan = \frac{lebar}{panjang}$$

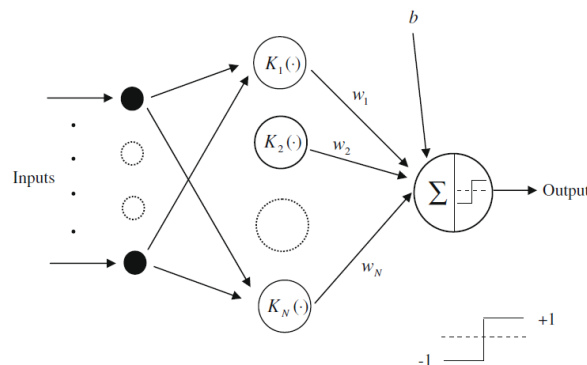
Panjang adalah panjang objek dan lebar adalah lebar objek. Fitur ini terkadang disebut rasio aspek (Ruuska et al., 2018).



Gambar 2. Gambaran Klasifikasi

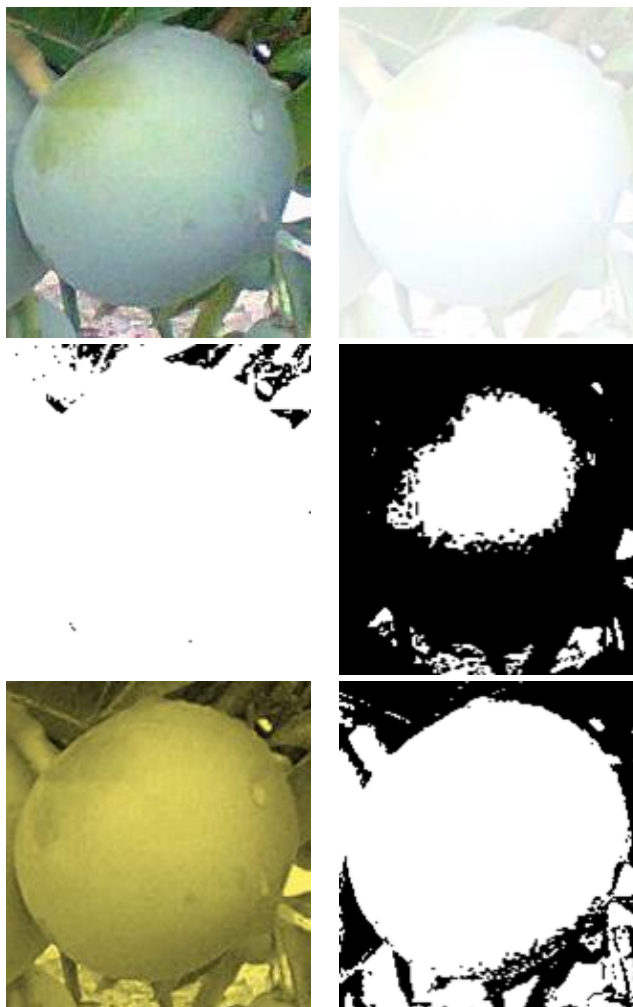
### Hasil dan Pembahasan

Pada penelitian ini dipilih gambar yang mempunyai kerumitan tinggi untuk menguji performa sekaligus akurasi metode *K - Means*. Metode *K - Means* menempuh beberapa tahap dalam proses klasifikasi teks, yaitu tahap membuat struktur transformasi warna secara spesifik, dan tahap klasifikasi, pada tahap pelatihan dilakukan proses analisis terhadap sampel gambar buah mangga.



Gambar 3. Hasil Clusterin Pixel Buah Mangga

Proses klasifikasi buah mangga apel menggunakan hasil ekstraksi ciri data pengujian dan hasil proses klasifikasi pelatihan. Hasil proses ini berupa nilai indeks citra dari fungsi keputusan yang terbesar yang menyatakan kelas dari data pengujian. Kelas yang dihasilkan dari proses klasifikasi pengujian sama dengan kelas data pengujian, maka pengenalan citra menggunakan metode kmeans klustering dinyatakan benar. Berikut gambar hasil treshld citra.



Gambar 6. Hasil Segementasi

Proses uji coba dengan menggunakan LS-SVM dilakukan pada data training dan pada data uji. Proses uji coba pada data training dilakukan untuk mendapatkan nilai gamma dan sigma yang paling optimal. Hasil dari proses uji coba pada data training dilakukan berkali-kali sehingga mendapatkan nilai akurasi 100%.

Tabel 4. Konfusi Matriks

Aktual	Klasifikasi	
	Positif	Negatif
Positif	(59 TP)	(1 FN)
Negatif	(1 FP)	(59 TN)

Waktu Komputasi yang dibutuhkan untuk proses ekstraksi fitur untuk setiap dataset adalah rata-rata sebesar 0,85 detik, sedangkan waktu komputasi untuk uji coba

pada data uji adalah rata-rata sebesar 0,006 detik. Dapat disimpulkan bahwa waktu komputasi untuk setiap dataset adalah rata-rata sebesar 0,856 detik.

### **Kesimpulan**

Hasil penelitian klasifikasi kultivar buah mangga dan apel menggunakan metode pengolahan citra cluster k-means dengan metode backpropagation dapat disimpulkan sebagai berikut: Dari percobaan yang sudah dilakukan penggunaan 1 hidden layer lebih efektif dengan akurasi tertinggi dan waktu tercepat dari pada menggunakan 2 hidden layer. Dari percobaan dengan 2 model jaringan syaraf tiruan, trainlm merupakan model yang sangat baik dalam hal waktu dan akurasi jika menggunakan 1 hidden layer. Nilai Rata-rata hasil akurasi tertinggi dalam proses klasifikasi jenis Buah mangga ini menggunakan epoch 5000, goal 0.0001 dan learning rate 0.1 ialah hidden layer 1 dengan model trainlm jumlah rata-rata akurasi sebesar 89.95% dengan rata-rata waktu 6.2 detik. Sedangkan hidden layer 2 pada model traingdx diperoleh rata-rata akurasi 89.95% dengan waktu 19 detik.



## Daftar Pustaka

- Abdul, K., & Adhi, S. (2013). Pengolahan Citra Teori dan Aplikasi. *Penerbit Andi*.
- Alamuru, S., & Jain, S. (2021). Video event classification using KNN classifier with hybrid features. *Materials Today: Proceedings*, xxxx. <https://doi.org/10.1016/j.matpr.2021.03.154>
- Arslan, H., & Arslan, H. (2021). A new COVID-19 detection method from human genome sequences using CpG island features and KNN classifier. *Engineering Science and Technology, an International Journal*, 24(4), 839–847. <https://doi.org/10.1016/j.jestch.2020.12.026>
- Aulianida, D., Liestyasari, S. I., & Ch, S. R. (2019). PENERAPAN METODE FUZZY TOPSIS DAN FUZZY SAW DALAM MENENTUKAN LOKASI WISATA DI NUSA PENIDA. *Journal of Chemical Information and Modeling*, 53(9), 1689–1699.
- Malang, N. (2016). *Klasifikasi Jenis Mangga Berdasarkan Bentuk Dan Tekstur Daun Menggunakan Metode K-Nearest Neighbor*. 8, 247–253.
- Narendra, N., Ramakrishnan, V., Thomas, A., Arvind Lambodhar, U., & Suhas, K. (2021). Characterization and feasibility of biogas yield using mango fruit peels and durva grass. *Materials Today: Proceedings*, 47(xxxx), 4624–4629. <https://doi.org/10.1016/j.matpr.2021.05.479>
- Nurmalasari, :, & Pratama, A. A. (2018). Sistem Pendukung Keputusan Pemilihan Supplier Menggunakan Metode Analytical Hierarchy Process (AHP) Pada PT Transcoal Pacific Jakarta. *Jurnal Teknik Komputer*, IV(2), 48–55. <https://doi.org/10.31294/jtk.v4i2.3509>
- Parida, P., & Bhoi, N. (2017). 2-D Gabor filter based transition region extraction and morphological operation for image segmentation. *Computers and Electrical Engineering*, 62, 119–134. <https://doi.org/10.1016/j.compeleceng.2016.10.019>
- Ristian, U. (2017). *Optimasi ekstraksi ciri kontur daun menggunakan interpolasi nearest neighbor dan b-spline uray ristian*.
- Ruuska, S., Hämäläinen, W., Kajava, S., Mughal, M., Matilainen, P., & Mononen, J. (2018). Evaluation of the confusion matrix method in the validation of an automated system for measuring feeding behaviour of cattle. *Behavioural Processes*, 148, 56–62. <https://doi.org/10.1016/j.beproc.2018.01.004>
- Sari, I. P., Hidayat, B., & Atmaja, R. D. (2016). *Perancangan dan Simulasi Deteksi Penyakit Tanaman Jagung Berbasis Pengolahan Citra Digital Menggunakan Metode Color Moments dan GLCM*. 215–220.
- Shokrzade, A., Ramezani, M., Akhlaghian Tab, F., & Abdulla Mohammad, M. (2021). A novel extreme learning machine based kNN classification method for dealing with big data. *Expert Systems with Applications*, 183(December 2019), 115293. <https://doi.org/10.1016/j.eswa.2021.115293>

- Singh, V., & Misra, A. K. (2017). Detection of plant leaf diseases using image segmentation and soft computing techniques. *Information Processing in Agriculture*, 4(1), 41–49. <https://doi.org/10.1016/j.inpa.2016.10.005>
- Sumaiya, M. N., & Shantha Selva Kumari, R. (2017). Gabor filter based change detection in SAR images by KI thresholding. *Optik*, 130, 114–122. <https://doi.org/10.1016/j.ijleo.2016.11.040>
- Utami, A. T. R. I., Informatika, P. S., Komunikasi, F., Informatika, D. A. N., & Surakarta, U. M. (2017). *Implementasi metode otsu thresholding untuk segmentasi citra daun*.
- Younesi, A., & Amirani, M. C. (2017). Gabor Filter and Texture based Features for Palmprint Recognition. *Procedia Computer Science*, 108(June), 2488–2495. <https://doi.org/10.1016/j.procs.2017.05.157>